Big Data and Extreme Computing Applications Breakout Report

BDEC workshop, July 15th 2015, Frankfurt

David Keyes (KAUST)
Rosa M Badia (BSC)

Breakout participants January 2015

- James Ahrens
- Jean-Claude André
- Rosa Badia
- Thierry Bidot
- Mathis Bode
- Fernando Cucchietti
- Anshu Dubey
- Françoise Genova
- David Keyes
- Jysoo Lee
- Takemasa Miyoshi
- Morris Riedel
- Alejandro Ribés
- Joel Saltz
- Alex Szalay
- Bill Tang
- Akira Ukawa
- Pier Luigi Vidale

Profile of the group

- Nations represented
  - Asia: Japan, Korea
  - Americas: USA
  - EU: France, Germany, Iceland, Spain, UK
  - Middle East: Saudi Arabia
- Disciplines represented
  - PDE-based: combustion, climate/meteorology, fusion
  - Observation-based: astronomy, biomedical informatics
  - Tools-based: PDE frameworks, runtime systems

Comparing "HPC" and "BDA"

- **The group preferred to reject the given labels, considering instead "numerically intensive" (NI) and "data intensive" (DI), both of which can avail themselves of HPC.**
  - **Disclaimer: These are also imperfect; the right labels is a "work in progress."**
- **Both NI and DI approaches share the common challenge of gaining scientific insights, making prediction, and quantifying uncertainty**
  - **NI through first principles models**
  - **DI through statistical models**

Application types (1/2)

- Third paradigm
  - PDE-based models
  - Particle-based models
  - Linear algebra-based models (e.g., DFT)
  - Image processing
- Fourth paradigm
  - Archiving and retrieving from massive data sets
  - Clustering
  - Searching
  - Knowledge discovery

Application types (2/2)

- Combinations of Third and Fourth paradigms
  - Fourth informs Third
    - Inverse problems
    - Data assimilation
    - Visualization and computational steering
  - Third informs Fourth
    - Design of experiments
  - Both paradigms in a virtuous loop

How to distinguish?

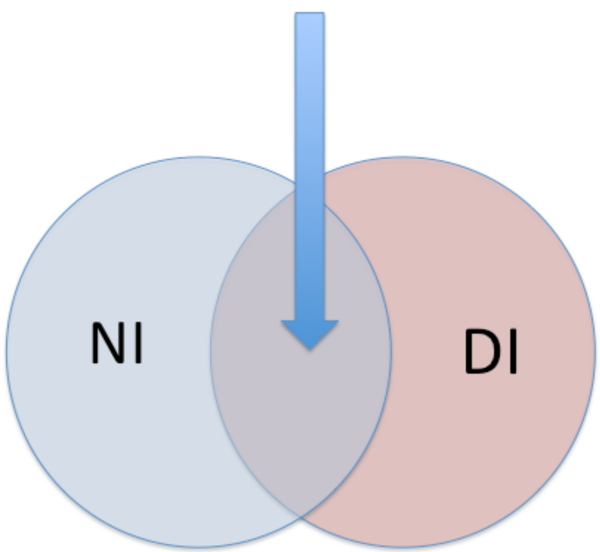
- “Amdahl number” (Gray & Szalay),
  - Ratio of bits by CPU cycle/s: O(10^-3) for NI and O(1) DI
- Data access pattern is a useful distinguishing feature

Questions to the breakout

- **What are the main differences and commonalities between the HPC and BDA requirements/technologies/working-assumptions in this area?**
- **Are there common needs/problems/interfaces could serve as the basis (or as stepping stones) along a path to (some reasonable level of) infrastructure and application convergence?**
- **Are there inter-domain testbeds that combine “BDA” and “HPC” workflows in ways that could help uncover pathways toward convergence?**
- **What is/are the technology or new research that may be a game changer?**
- **What action would be your number one priority to be taken rapidly to ensure success of the convergence of Extreme computing and Big Data infrastructures?**

Main differences and commonalities?

- There is a growing overlap between NI and DI in applications. Many NI applications produce Big Data and DI is a growing consumer of HPC capabilities.
- There is a growing opportunity for common approaches, common software frameworks, and even reuse of data structures.



Significant overlap in the spatio-temporal data realm

Main differences and commonalities?

- NI has always been oriented towards causality (dynamics, hypothesis-driven), whereas DI has traditionally been oriented towards archiving data for future discovery
  - "The most important thing about Big Data is that you can go back to it." (Amazon)
- NI data, arising from continuous models on meshes or swarms, tends to be highly structured and highly correlated, which may not be true of DI data
- NI data tends to do lots of writes, whereas DI tends to read many times and write few times.

Main differences and commonalities?

- I/O - Storage:
  – DI has a tendency to use machines with larger memory spaces and I/O infrastructure more oriented towards random access
  – NI usage is dominated by bulk streaming writes
  – Still, a growing degree of parallelism with asynchronous, distributed computation is a trend in applications from both sides
- In NI applications the main action occurs in *memory* and I/O tends to be *regular*, whereas in DI the I/O stream is much more *complex*.

Axis	Sub-axis	Numerically Intensive	Data Intensive
Hardware	Nodes and Interconnect	High performance and power	Lower performance and power
	Storage	Separate, independent	Integrated
SW	Synchronization	Tightly coupled	Loosely coupled
	Reliability	Checkpoint restart	Replication
Workload	Number of Users	**Single per node**	**Multiple per node**
	Data	Dynamic, heterogeneous (unstructured grid)	Static, homogeneous (text, images)
	Algorithms	Global	Distributed
	User Interface	**Complex Application**	**Simple Web**
	Data Model	**Files**	**Database**
Workflow	Scheduling	Batch	Interactive
	Analysis	Offline post-processing	Online
	I/O	Bulk parallel writes	Streaming writes

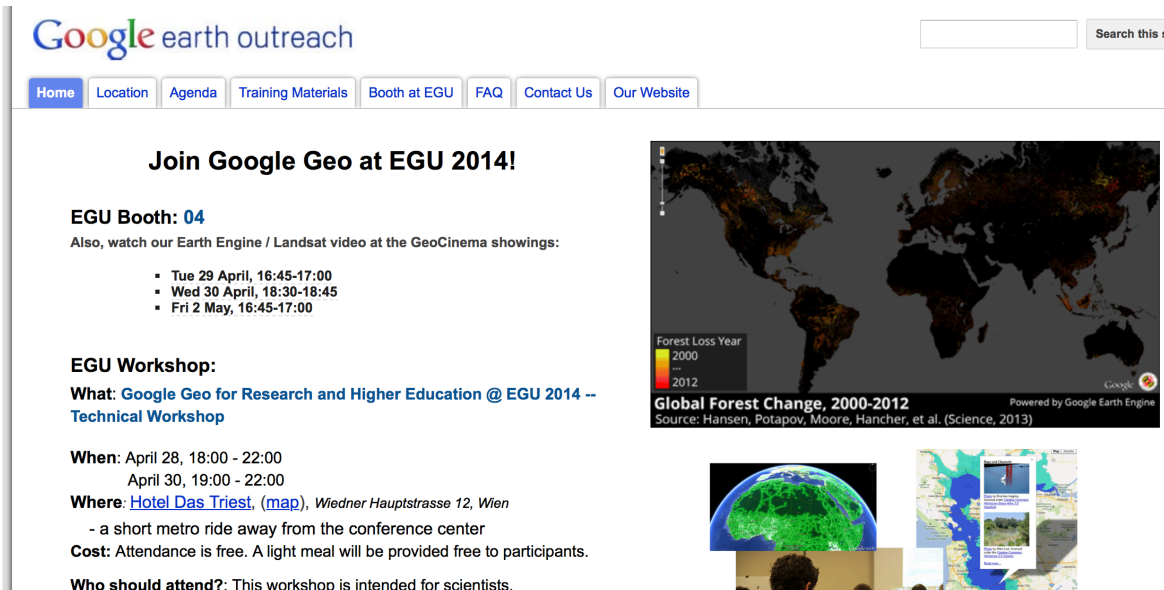
c/o Ahrens, 2012

Common needs/problems/interfaces that could serve for application convergence?

- To some degree, convergence has been dictated by the *hardware* trends and what vendors make available
  - Less convergence in *software*.
- Under all hardware scenarios, data movement is becoming relatively increasingly expensive and analytics should be computed *in situ,* or as close as possible to the data source.
  - This may dictate changes in the store vs. recompute spectrum.

Testbeds for convergence

- Joint initiative between academia and industry to investigate applicability of high end "Google-type" platforms to scientific, medical, and engineering problems
  - e.g., Google Earth collaboration with European Geoscience Users
  - https://sites.google.com/a/earthoutreach.org/google-egu-2014/



Game changers?

- Combination of embedded processing components, 3D memory technology, and out-of-core memory algorithms exploiting burst buffers
  - Have the potential of enhancing *massive data processing* while in transit and therefore enable many of the computations needed to extract knowledge in *both domains*
- No clear answers of what can be the game changers:
  - Systematic assessment is needed to determine what are the *true requirements* needed by users/scientists
  - There is a need to investigate models that *get away from classical batch processing* and take more into account how the information is rendered accessible to decision makers

Highest priority for "convergence"?

- Build open data repositories and develop data challenges that highlight the unification of the techniques
  - Ranging from data services to facilitate access, down to general data query and processing for investigation and exploration.
- Conferences and workshops that foster a unification of the research communities
  - Now highly fragmented, as evidenced by several presentations at the workshop.
- Friendlier allocation processes for major facilities for data intensive workflows
  - Currently adapted to the needs of a typical simulation user (e.g., batch, temporary use of storage)
  - Allow the persistent storage of relevant data bases

Highest priority for emergence?

- Algorithmic research to achieve scaling of parallel techniques on the Big Data front
- New hierarchical representations of data
  - Instead of storing all data at finest resolution store representative sets at finest resolution, along with recursively space and time coarsened sets, and statistics.
- New data structures
  - e.g., based on space filling curves which may be efficient in accessing data at adaptively increasing scales
  - Coarser scales useful for transmission and low-resolution visualization
  - Finer scales useful for analysis once the phenomena of interest are identified.

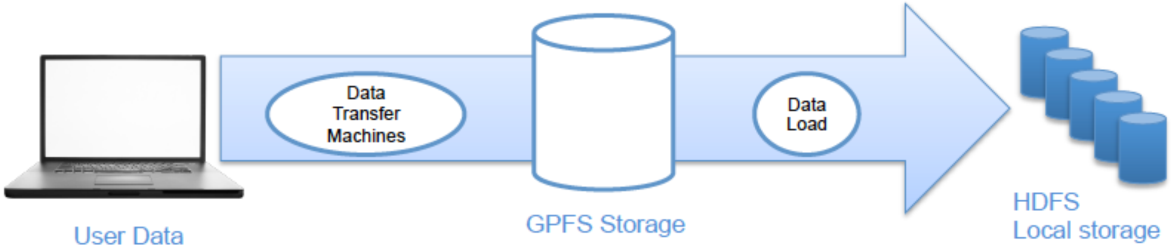
Highest priority for emergence?

- Requirements on the SW/HW environment
  - Portability
    - And performance portability
  - Usability
  - *In situ* use of the data
    - On the fly extraction
  - Means of tracking provenance through data transformations

An example of use of SC for BigData

spark4MN

- Spark deployed in MareNostrum supercomputer
  - Memory (connected IB) 94 TB
  - Number of Cores 48,448
  - Distributed Storage 1,5 PB
- Set of commands and templates
  - Spark4mn
    - Sets up the cluster, and launches applications, everything as one job.
  - spark4mn_benchmark
    - N jobs
  - spark4mn_plot
    - Metrics



Conclusions

- We should not force a "shotgun" marriage of "convergence"
  - When a love-based marriage is inevitable in the near future
- The ultimate test of our efforts will be whether the emergence of NI+DI applications meets *mission-critical needs in scientific discovery and engineering design*
- We need a follow-up meeting with additional input from extreme computing big data people

THANKS!